2010 SREE Conference Abstract Template

Abstract Title Page

Title:

Instructional Interactions of Kindergarten Mathematics Classrooms: Validating a Direct Observation Instrument

Author(s):

Christian Doabler, **Pacific Institutes for Research**; Keith Smolkowski, **Oregon Research Institute**; Hank Fien, **University of Oregon**; Derek B. Kosty, **Oregon Research Institute**; Mari Strand Cary, **Pacific Institutes for Research**

Abstract Body

Background/Context:

Over the course of the past three decades, the measurement of teaching processes has been a cornerstone of education research (Shavelson, Webb, & Burstein, 1986). Results from the study of instruction have played an influential role in generating a knowledge base of effective pedagogical practice (Brophy & Good, 1986). Recently, researchers have begun to investigate the quality of instruction with a specific focus on the instructional interactions that occur between teachers and their students (Pianta & Hamre, 2009; Smolkowski & Gunn, 2009).

One method for collecting information about the quality of instruction and student learning is direct observation (Ball & Rowan, 2004; Kennedy, 1999). Direct observation is a sensitive method for handling different sources of variance such as time, observer effects, and moment-to-moment changes of teacher and student behavior (Snyder et al., 2006). Researchers can use direct measures of observed instruction to investigate mediating variables and explore the relationships between student outcomes and teaching practices (Brownell et al., 2009; Hiebert & Grouws, 2007; Vaughn & Briggs, 2003).

Interestingly, few studies have used direct observation to investigate the quality of instruction provided in kindergarten mathematics classrooms. Because a successful start in kindergarten is critical for subsequent achievement growth in mathematics (Bodvoski & Farkas, 2007; Chard et al., 2008; Morgan, Farkas, & Wu, 2009), there is need to identify and measure the effective teaching practices that facilitate learning during this critical period of children's early education. We reason direct observation is the most applicable approach for identifying the variables that mediate the impact of kindergarten instruction. This paper reports on our use of a direct observation instrument in kindergarten mathematics classrooms.

Purpose:

In this paper, we report research focused directly on the validation of the Coding of Academic Teacher-Student interactions (CATS) direct observation instrument. We are using classroom information gathered by the CATS instrument to better understand the potential mediating variables hypothesized to influence student achievement. Our study's purpose is to gather the kinds of validity evidences that match the proposed interpretations and uses of the CATS instrument (Kane, 2008; Messick, 1995). Therefore, we first explore the content aspect of construct validity by collecting information about the content relevance and representativeness of the observation instrument (Messick, 1995). Second, concerned about the consistency of the observation data collected across a number of independent observers, we measure inter-observer agreement. Finally, we focus on the criterion-predictive aspect of construct validity and investigate the relationship between student gains and the quantity of teaching. We anticipate that classrooms with higher rates of practice opportunities and teacher demonstrations will foster greater increase in classroom-level achievement than classrooms with fewer instructional interactions. To address this assumption, we explore curriculum based measurement (CBM) scores and information of observed teaching practice.

Setting:

We conducted the current research within the context of a federally funded randomized control efficacy trial, the Early Learning in Mathematics: Efficacy Trials in Kindergarten

Classrooms (ELM-ETKC) study (Baker, Chard, Clarke, & Smolkowski, 2008). We are investigating the efficacy of the Early Learning in Mathematics (ELM) kindergarten curriculum in two states over the course of four years. This paper concerns Year 1 of the study, which involved 65 general education kindergarten classrooms in rural and urban schools across the state of Oregon. Participating classrooms were located within high-poverty Title 1 schools. Thirty-seven classrooms provided a half-day kindergarten program and twenty-eight offered a full-day program. Average class size was 19.3 in the ELM classrooms and 19.4 for comparison classrooms.

Participants:

A total of 66 teachers and 1,450 kindergarten students participated in the first year of the ELM-ETKC study. Of the 66 teachers, 97% were female. Participating teachers reported an average of 9.69 years of teaching experience, and a mean of 6.25 years teaching at the kindergarten level. For each condition, we plan to report student demographic data such as percentages for ethnicity, special education services, limited English proficiency, and free or reduced lunch.

Intervention:

We conducted the current research within the ELM-ETKC randomized control efficacy trial (Baker et al., 2008). The development of the ELM intervention focused on curriculum content and design, and teachers' instructional interactions with students. However, this present research focuses primarily on the CATS observation instrument used across all classrooms, including those that did not use the ELM curriculum.

The CATS instrument systematically measures the instructional interactions that occur between teachers and students during kindergarten mathematics instruction (see Figure 1 in Appendix B for a copy of the tool). We ground the CATS instrument in the best evidence on effective instruction for early mathematics learning (Gersten et al., 2008, 2009; Jayanthi, 2008; National Mathematics Advisory Panel, 2008). Certain features of the tool were also adapted from the Student-Teacher Interactions Context Observation instrument (Smolkowksi & Gunn, 2009).

The CATS instrument includes sections for documenting information about the context of the observation and the instructional interactions hypothesized to influence mathematics achievement. The context section entails three components (a) instructional start and finish times for the math activity, (b) type of content targeted in the activity, and (c) type of instructional format (small-group or whole-class). The mathematical content areas include (a) number and operations, (b) geometry, and (c) measurement.

The instructional interaction section focuses on six behaviors: (a) teacher models, (b) group responses, (c) individual responses, (d) covert responses, (e) student mistakes, and (f) teacher-provided academic feedback. Observers code behavior occurrences in a continual, serial fashion. Consequently, we are able to consider the total number of observed behaviors and the rate in which they occurred.

In addition, the data will also allow us to consider the frequency and quality of instructional interactions that occur between a teacher and their students. Features of an instructional interaction include a sequence of explicit teacher modeling and academic feedback, and systematic practice opportunities for students. There is a growing body of evidence that suggests these instructional interactions are an integral part of effective classroom instruction (Chard et al., 2008; Gersten et al., 2008, 2009; Jayanthi, 2008; Smolkowksi & Gunn, 2009).

An example of an instructional interaction might consist of a teacher demonstrating the physical attributes of a cube and then asking a group of students to count the sides of the shape, followed by a request for an individual student to state the shape's name. To extend student understanding, the teacher would provide academic feedback immediately following each response. Using this example, observers would code the following sequence on the CATS observation instrument: teacher model, group response, academic feedback, individual response, and academic feedback.

Research Design:

To address the efficacy of the ELM curriculum, the ELM-ETKC project conducted a randomized controlled trial utilizing a randomized block design, with classrooms assigned within schools to treatment and control conditions. Random assignment within schools resulted in 35 treatment (ELM) classrooms and 30 comparison ("business as usual") classrooms. Instruction in the treatment condition consisted of the ELM curriculum. Teachers in comparison classrooms used a variety of commercially available and teacher-developed materials.

Data Collection and Analysis:

Currently, all necessary data have been collected, but the analyses have not been fully completed. We thus provide information on the data collected, the analyses conducted to date, and the additional analyses planned for the presentation.

Project staff collected observation data at three points within each classroom during the school year. Previous work documented the instrument's reliability and capacity to predict important outcomes when collected during reading instruction (Smolkowski & Gunn, 2009). Preliminary analyses from the present study have also demonstrated that observers can collect consistent data across treatment conditions. From our prior work, we do not expect growth in the rate of observed student-teacher instructional interactions, but we will nonetheless test this assumption. If we find no systematic growth, we will examine the predictive validity of teachers' aggregate rate of student practice across the school year (i.e., average of the three observations). If we find that the slope differs significantly from zero, we will use the growth parameters to predict growth in student outcomes.

The project also provides student level CBM scores collected at five time points across the past school year. Prior research has shown the CBM measures to be psychometrically valid and predictive of comprehensive measures of mathematics (Clarke, Baker, Smolkowski, & Chard, 2008). These measures will serve as the dependent variables for the prediction models.

Based on the nested structure of the data, with students within classrooms, the remaining analyses will use hierarchical linear models to test the predictive validity of observed opportunities for student practice. We will first estimate student growth in CBM scores across the school year and then predict the growth parameters with, depending on the presence of growth in teacher behaviors, either classroom-level (aggregate) opportunities for student practice or the growth parameters thereof.

Findings:

For our first objective, we administered a web-based survey to assess the content validity of the six observed behaviors captured by the instrument. The target population for the survey consisted of experts with extensive knowledge in the areas of elementary mathematics curricula and instruction. Of the 12 experts contacted, 7 (58%) completed the survey. The survey comprises 12 items, of which six obtain information about content relevance and six obtain information about content representativeness (Messick, 1995). Items on the survey used a 4-point scale, with 1 being the lowest and 4 the highest. We considered a score of 3 as a minimally acceptable score for all items.

With respect to content relevance, teacher model and academic feedback both received an average rating of 4 on the content relevance items. Of the four student behaviors, student mistake and individual response received an average rating of 4. The covert response and group response behaviors earned average ratings of 3.33 (*SD* =.52) and 3.83 (*SD* =.42), respectively.

In regard to content representativeness (Messick, 1995), all six items met the minimum mean rating of 3. In fact, teacher model, academic feedback, individual response, and student mistake received average ratings of 4. Group response received a mean rating of 3.83 (*SD* =.14), while covert response earned an average of 3.67 (*SD* =.51).

For our second objective, we used a reliability index of inter-observer agreement to measure the consistency of data collected across a number of independent observers. Based on 46 paired observations, average inter-observer agreement was 90% for student practice opportunities, 84% for teacher behaviors, and 90% across all instructional interactions.

For the study's third objective, we plan to investigate the relationship between student gains and the quantity of teaching. Preliminary, non-nested analyses *suggest* strong relationships between student practice opportunities and change in CBM scores. We expect similar results from the multilevel models described above, and have found such relationships from previous research on literacy instruction (Smolkowski & Gunn, 2009). Replication of these findings for the teaching of mathematics would demonstrate the generalizability of the student-teacher instructional interactions across instructional domains. Further analyses are currently underway.

Conclusions and Implications:

Results from the current study of mathematics instruction in kindergarten could contribute to the accumulating knowledge base of effective pedagogical practice. Preliminary findings indicate that the CATS tool is reliable and valid for documenting the quantity and quality of student-teacher interactions in kindergarten mathematics classrooms. This study has several implications for the observation of classroom instruction. First, the recent advent of Response to Intervention calls for the delivery of high quality instruction to occur in the general education classroom across content areas. School personnel such as instructional coaches and school psychologists will need reliable and valid tools for determining whether students are receiving highly interactive math instruction in content that aligns with state standards. The CATS tool could potentially serve as such a mechanism. Additionally, generating estimates of rates of student math practice opportunities that are associated with strong student mathematics outcomes could be useful for the design and development of new mathematics curricula. Likewise, patterns of teacher models and academic feedback that appear to associate with strong student math outcomes could have implications for teacher professional development of early mathematics instruction.

Appendixes

Appendix A. References

Baker, S. K., Chard, D. J., Clarke, B. S., & Smolkowski, K. (2008). *Early learning in mathematics: Efficacy trials in kindergarten classrooms.* Research funded by the U.S. Department of Education, Institute of Education Sciences (CFDA 84.305A). Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Ball, D. L., & Rowan, B. (2004). Introduction: Measuring instruction. *Elementary School Journal, 105,* 3-10.

Bodovski, K. & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *The Elementary School Journal, 108,* 115-130.

Brownell, M. T., Bishop, A. G., Gersten, R., Klingner, J. K., Penfield, R. D., Dimino, J. et al. (2009). The role of domain expertise in beginning special education teacher quality. *Exceptional Children, 75,* 391-411.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. C. Wittrock (Ed.), *Handbook of research on teaching* (3rd ed., pp. 328-375). New York: Macmillan.

Chard, D. J., Baker, S. K., Clarke, B., Jungjohann, K., Davis, K., & Smolkowski, K. (2008). Preventing early mathematics difficulties: The feasibility of a rigorous kindergarten mathematics curriculum. *Learning Disability Quarterly, 31*, 11-20.

Clarke, B., Baker, S., Smolkowski, K., & Chard, D. J. (2008). An analysis of early-numeracy curriculum-based measurement: Examining the role of growth in student outcomes. *Remedial and Special Education, 29,* 46-57.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to Intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Gersten, R., Chard, D. J., Jaynthi, M., Baker, S., Morphy, P., & Flojo, J. (2008). Mathematics instruction for students with learning disabilities or difficulty learning mathematics: A synthesis of the intervention research. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester, Jr. (Ed.), *Second handbook of research on mathematics teaching and learning: Vol. 1. A project of the national council of teachers of mathematics* (pp. 371-404). Charlotte, NC: Information Age Publishing.

Jayanthi, M., Gersten, R., & Baker, S. (2008). Mathematics instruction for students with learning disabilities or difficulty learning mathematics: A guide for teachers. Portsmouth, NH: RMC Research Corporation, Center on Instruction.

Kane, M. T. (2008). Terminology, emphasis, and utility in validation. *Educational Researcher, 37,* 76-82.

Kennedy, M. M. (1999). Approximations of indicators of student outcomes. *Educational Evaluation and Policy Analysis, 21,* 345-363.

Messick, S. (1995). Standards of validity and the validity of standards in performance assessments. *Educational Measurement: Issues and Practice, 14*(4), 5-8.

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities, 42,* 306-321.

National Mathematics Advisory Panel, (2008). *The final report of the national mathematics advisory panel.* Retrieved March 17, 2008, http://www.ed.gov/about/bdscomm/list/mathpanel/pre-report.pdf

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher, 38*, 109-119.

Shavelson, R. J., Webb, N. M., & Burstein, L. (1986). Measurement of teaching. In M. C. Wittrock (Ed.). *Handbook of research on teaching* (3rd ed., pp. 50-91). New York: Macmillan.

Smolkowski, K., & Gunn, B. (2009). *The reliability and validity of student-teacher interaction and context observations collected during kindergarten reading instruction.* Manuscript submitted for publication.

Snyder, J., Reid, J., Stoolmiller, M., Howe, G., Brown, H., Dagne, G., Cross, W. (2006). The role of behavior observation in measurement systems for randomized prevention trials. *Prevention Science*, *7*(1), 43-56.

Vaughn, S. R., & Briggs, K. L. (Eds.) (2003). *Reading in the classroom: Systems for observing teaching and learning*. Baltimore: Paul H. Brookes.

Appendix B. Measures

Figure 1. Example of CATS observation instrument